Imperial College London

Department of Earth Science and Engineering

MSc in Applied Computational Science and Engineering

Independent Research Project
Final Report

Prediction of wildfire duration with image-based Machine learning

by

Hansong Xiao

Email: hansong.xiao21@imperial.ac.uk
GitHub username: acse-hx221
Repository: https://github.com/ese-msc-2021/irp-hx221

Supervisors:

Dr. Sibo Cheng
Dr. Rossella Arcucci

August 2022

1 Abstract

Wildfire forecasting problems usually rely on complex grid based mathematical models, mostly involving CFD and Celluar Automata, but these methods have always been computationally expensive and difficult to deliver a fast decision pattern. In this paper, we provide machine learning based approaches that solve the problem of high computational effort and time consumption. This paper predicts the burning duration of a known wildfire by RF(random forest), KNN, and XGBoost regression models and also image based, like CNN and Encoder. Model inputs are based on the map of landscape features provided by satellites and the corresponding historical fire data in this area. This model is trained by happened fire data and landform feature map and tested with most recent real value in the same area. By processing the input differently to obtain the optimal outcome, the system is able to make fast and relatively accurate future predictions based on landscape images of known fires.

2 Introduction

As humans evolved, fire seems to become more controllable. But wildfires are not included, wildfires are a natural phenomenon that is difficult to predict and occur with high frequency, more than 4000 fires happened in CA, USA since 2003 to 2016 (ANDELA et al. 2019). More seriously, wildfire can even be called natural hazard (McCaffrey 2004). Wildfires endanger most of the world(Bond & Keeley 2005), 4 million km^2 to 6 million km^2 (depends on different estimate method) of land are affected by wildfire every year(Youssouf et al. 2014), which is almost half the area of Europe.

Wildfire has huge negative impact on earth ecosystem, risking world natural system and health, hydrogeomorphic effects from wildfire jeopardize the quality of daily water in both human society and ecosystem(Robinne et al. 2018). Wildfire also threatens the world economy, house price of northwest Montana, a potential resort area in USA, was depreciated by unpredictable long-term wildfire(Stetler et al. 2010). It also results in serious injury or death for animals, a huge group of sheep burned in a wildfire in 2006(Madigan et al. 2011). To prevent multiple hazards, researchers have begun to engage in developing with wildfire prediction. Cellular Automata (Chopard & Droz 1998) and CFD (Computational Fluid Dynamic) (Anderson & Wendt 1995) were widely used. These traditional methods normally based on building a complex mathematical model that take all variables and interaction between features into consideration. One of the most important and pioneering model is Rothermal's equation(Rothermel 1972), but most of these models are hard to understand and lack of interpretability. Basically, solution to grid-based method is solving dynamic function or a system of partial differential equation(Alexandridis et al. 2008), which is notoriously time-consuming. Thus, grid-based methods were intensive on computation and required resources that are nearly impossible to achieve a fast-decision pattern. With the development of artificial intelligence, machine learning becomes a new research direction(Fradkov 2020) on wildfire prediction. Satellite imagery was used to help analyze geomorphic features to build models for better forecasting. However, it is challenging to develop a general and effective ML model in world-wide range for forecasting wildfire, due to the complicity of environmental parameters in different area.

Currently, most of ML model put insight on specific direction, like occurrence of human-caused wildfire(Rodrigues & de la Riva 2014), forecasting sensitivity map of wildfires(Ghorbanzadeh et al. 2019) and predicting likelihood of wildfires (Parisien et al. 2012). Most of the current research is focusing on possible future wildfire, which requires management group put energy on every possible location. 80 percent fire last no more than one day based on our extracted dataset (ANDELA et al. 2019), indicating fire management group are unable to respond effectively in a limited time. It would be more wise to keep eyes on preventing potential large fire. In this paper, we aims to deliver a fast-decision model that focus on predicting the likely extent of ongoing fires. ML model were split into two part, simple regression model and image-based model. In ML regression part, three methods were applied, including RF(Breiman 2001), KNN(Altman 1992), and XGBoost(Chen & Guestrin 2016)

to predict fire duration. In image-based part, two CNN based models were introduced, including multi-layer CNN and Encoder-Based CNN.

3 Data curation

3.1 Data Gathering

We train and validate our model using most of the fires that occurred in 13 years in a specific region, as shown in figure 1, ranging from 32.7N to 42N, -124.26W to -93.5W, which is mainly the Midwestern region of the United States, including California, Phoenix, Utah, Colorado, New Mexico, Kansas, and Texas. This part was chosen due to the similarity of geomorphology among them (Kandakji et al. 2020), the increasing number of wildfires with a expanded rate of burned area at $355km^2$ per year (Dennison et al. 2014), and an enormous economic loss excesses billions dollar per year (Boesch et al. 2000). If we consider the entire U.S., the geomorphic disparity between the West Coast and the East and sparsity of wildfire on East would impact the accuracy of the model.

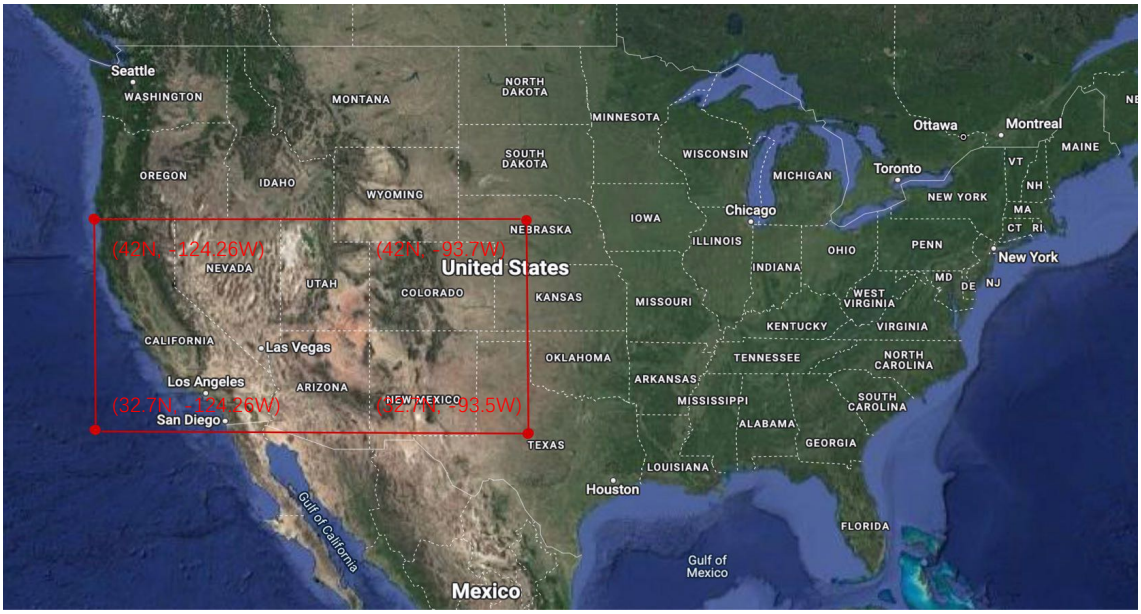


Figure 1: Target area

We decide use *altas fire dataset*(ANDELA et al. 2019) for fire data, which includes the coordinates, start and end times, burned area, perimeter, and duration of the fires. We selected fire duration as the label value, considering the duration becomes more significant in evaluation wildfire(Radke et al. 2019). Eighty percent of fires are small, which last less than one day. To avoid data imbalance, we randomly dropped half of small fire (duration less than 1 day).

The database was split into two parts, from 2003 to 2013 as a train set, and from 2013 to 2016 as a test set. The ratio of fire number between train and test is around 0.75, which is a reasonable number for data splitting. We expand each target coordinate in four directions for 30 kilometers range which is demonstrated in figure 2. The yellow point represents coordinate of a random fire, and each red line shows how the 30 kilometers radius of extension area. the square surrounded by green line is the feature map to gather. We use a 30km radius size to avoid repeat data extraction work, which is time-consuming if smaller radius were taken into consideration. Currently, three global scale databases were involved in this experiment, density map of above(or below) ground living biomass storage (Spawn et al. 2020), landform slope map (Theobald et al. 2015), land cover map of multiple surface characteristics (e.g. grass coverage, tree coverage, snow converge) (Buchhorn et al. 2020), and climate data corresponding to specific location range from 1979 to 2020 (e.g. precipitation, wind

speed) (Hersbach et al. 2018). In general, five input variables, tree cover, ${}^{r}\mathbf{X}_{z}^{tc} = \{{}^{r}x_{i,j}^{tc}\}$, geo-slope, ${}^{r}\mathbf{X}_{z}^{slope} = \{{}^{r}x_{i,j}^{slope}\}$, grass cover, ${}^{r}\mathbf{X}_{z}^{gc} = \{{}^{r}x_{i,j}^{gc}\}$, wind in u direction, ${}^{r}\mathbf{X}_{z}^{windu} = \{{}^{r}x_{i,j}^{windu}\}$ and wind in v direction ${}^{r}\mathbf{X}_{z}^{windv} = \{{}^{r}x_{i,j}^{windv}\}$ were obtained to predicting ${}^{r}\mathbf{D}^{z}$, duration. Here, i, j, the size of matrix, for slope range from 0 to 200, grass cover and tree cover range from 0 to 700, z, and wind range from 0 to 4, fire index, r, the radius of extension. Detailed table of dataset used shown in table 1.

Dataset	Description	Resoulution	Sources	Reference	Image Size
Global Fire Atlas with Characteristics of Individual Fires	A Global Scale dataset contains all individual fires location, times, and duration etc.	None	MODIS	(ANDELA et al. 2019)	None
Global ALOS mTPI	A global scale dataset of landform detail.	270 Metres	ERGo	(Theobald et al. 2015)	(200, 200)
opernicus Global Land Cover Layers	A global scale dataset provide land cover map	100 Metres	GlobCover, LC-CCI, MODIS, PROBA-V	(Buchhorn et al. 2020)	(700, 700)
ERA5 Monthly Aggregates - Latest Climate Reanalysis	A global scale dataset provide climate parameters	27830 Meters	ECMWF	(Hersbach et al. 2018)	(4, 4)

Table 1: Dataset Detail

Figure 2: Example of fire extension

3.2 Data Standardization

Data standardization were applied for all inputs due to different scale of vegetation cover, normalizing all input in the same scale. *min-max* normalization (Jayalakshmi & Santhakumaran 2011)were adopted to resize all data into 0 to 1 range, as the equation shown below, where X indicates each input matrix, where f is the feature name, r representing radius, and z regrading fire index.

$$ {}^{r}X_{z}^{f} = \frac{{}^{r}X_{z}^{f} - min({}^{r}X_{z}^{f})}{max({}^{r}X_{z}^{f}) - min({}^{r}X_{z}^{f})} \tag{1} $$

4 Methodology

This section discusses several models designed for fast-decision fire prediction based on different algorithms. To ensure the consistency of result, input and output are the same for models on different basis, three ML regression models maintain the same input and image-based ML model comparing with the same input, input contains five possible different images, five feature maps, which are grass cover, tree cover, landform slope, windu and windv, output is fire duration. We make use of three datasets, data from 2003 to 2013 in target area regarded as train data, from 2013 to 2016 classified as test data.

The prediction result is hard to evaluate by RMSE, because the severity range of fire is not precise. In numerical aspect, it would be a prediction with more than 50 percent error if a test data has a predicted value 20 but with a label value 30. However, in aspect of fire prevention, both 20 and 30 indicate this fire exists for a long term, showing the seriousness of this fire, thus it could be defined as a reasonable prediction, which can be classified as the true prediction. To balance the ambiguous prediction result from numerical and prevention aspect, another measurement was introduced to evaluate all the possible outcomes, which is called accuracy plot.

A prediction was regraded as TP ,true prediction, if the prediction are greater than critical value, otherwise it is FP (false prediction). Calculating the ratio of positive out of total prediction, as the equation shown in below equation 2, gives the accuracy of prediction. This ratio varies with different critical value, critical value with a range of 0 to maximum label value. Thus, an accuracy plot was obtained and the lowest point in the accuracy curve presented the worst situation of each method.

$$Accuracy = \frac{TP}{TP + FP} \tag{2}$$

where TP is the true prediction, which has the prediction greater than critical value, FP is false prediction, which has the prediction smaller than critical value.

Meanwhile, another equation were used to compute RMSE (relative MSE) of each singe prediction, which gives the ratio between predicted days and original fire duration.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(\frac{\left\| {}^{r}D_i^z - {}^{r}\hat{D}_i^z \right\|}{\left\| {}^{r}\hat{D}_i^z \right\|})} \tag{3}$$

where ${}^{r}D_i^z$ is the label value in radius r and fire index z, $\hat{y}$ is the prediction value.

4.1 Regression modeling

The Regression model aims to give a simple prediction. This model regards different standardized feature (vegetation cover, slope of topographic map, wind and precipitation)as input to predict fire severity. Output is a measurement variable that could represents the fire severity (fire size, fire duration), fire duration were used because it shows more accurate rather than other label value. For the radius, 5km was adopted here as the figure 8 shows that radius has nearly no effect on result.

4.1.1 Data preprocessing

We initialized the collected data and divided into different ranges of inputs. In order to explore the affection of different range of input, we cropped original image with a radius 30km to smaller radius, from the center point to corresponding scale and calculate average value of each image in different radius. Then, we obtained a series of average value with different radius. The average value is the

input for regression model. For example, the average value of 80 × 80 in the middle of original map will be used as input if we want a range of 3km.

$$ {}^{n}X_{f}^{r} = \frac{\sum x_{ij}}{i \times j} \tag{4} $$

where X is the average value of input, r is the radius, n is fire index and f is feature name. Below is a demonstration of crop process when radius equal to three kilometers.

$$ X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,600} \\ \vdots & \ddots & \vdots \\ x_{600,1} & \cdots & x_{600,600} \end{bmatrix} \overset{radius=3km}{\longrightarrow} \begin{bmatrix} x_{270,270} & \cdots & x_{270,330} \\ \vdots & \ddots & \vdots \\ x_{330,270} & \cdots & x_{330,330} \end{bmatrix} $$

4.1.2 Random Forest regression model

We use random forest (RF) (Breiman 2001) as first regression method. Random forest is an extension branch of bagging algorithm, and RF is based on a series of Decision Trees(DTs). Random Forest develops models, which consists of a series of trees, that predict individually and separately. The output integrates the results of each predictor by a majority rule to forecast an output or one classification. In regression study, different sets of sample were obtained from initial dataset, the result is the average value of all the trees. The hyperparameters contain the number of individual predictor *n*, the depth of each predictor *d*.

4.1.3 K-nearest neighbors algorithm model

K-nearest neighbors (kNN) is a supervised learning algorithm, working on both classifier and regression (Altman 1992). In classification, KNN assigns each single data point into the same category as the highest frequency of *k* nearest neighbors. In regression, calculating the average label value of *k* nearest point of the input train data to get the prediction value. In KNN model, input is the mean value of each feature map, and we keep use fire duration as output. The hyperparameter is the number of neighbors *k*.

4.1.4 XGBoost

Extreme Gradient Boosting (Chen & Guestrin 2016) is another ensemble learning algorithm derived from Gradient Boost Decision Tree (GBDT) (Ke et al. 2017). XGBoost take every input data into account as each input result one tree. XGBoost also consider the gradient of data, which saves calculation time. XGBoost gives regularized term for loss function, which penalties the complexity of model and compensate the learning term, thus XGBoost is more effective and more capable of preventing over-fitting. The researched hyperparameters are number of estimator *n*, learning rate *l*, and max depth *d*.

4.2 Image Based Modelling

This part aims to explore the performance of image based model in predicting fire severity. To ensure the accuracy of forecasting, we tried permutation of different input and decide vegetation cover and landform slope are final input as the give the best accuracy. The same output value or label value, fire duration, were selected to ensure the comparability between models. We make use of the image as input instead of the average value. All feature map were resized to the same and placed in different channels of one data array as the traditional CNN. Moreover, the fire duration dataset from 2003 to 2013 was used as train dataset and 2013 to 2016 as test dataset. 5km data radius were selected to avoid unnecessary computation and keep the same radius with regression ML model.

4.2.1 Data preprocessing

Different range of image are not interested in this model, since the radius presents little impact on final result. The data preprocessing method applied is resize, which scale all image to the same size by an interpolation method. Resize keeps the image the same but only change width and height. This approach calculates the value of a new image pixel based on the value of the original pixel points by bi-linear interpolation(Smith 1981) shown below.

$$x_{ij} = [(1-p)x_{ij} + px_{i+1,j}](1-q) + [(1-p)x_{i,j+1} + px_{i+1,j+1}]q \tag{5}$$

where p, q is the fraction of distance bewteen point, x is the pixel point, i, j is the index of input point.

4.2.2 Multilayer Convolutional Neural Network (CNN)

CNN, one of the most famous machines learning neural network(Albawi et al. 2017), extracted the significance feature from original data by mathematics operation convolution, which is used in this model for wildfire prediction. CNN has the ability to process large amounts of data due to the number of its neural, CNN could better recognize and learn for image type input. The structure of this model is shown in below figure 3. ReLu and linear was used as activation function. A 3×3 kernel size and 32 channel number are selected in convolutional layer to maximum keep the feature of image. Two dropout layer is added in this model in case of over-fitting. To avoid gradient vanishing (Mastromichalakis 2020), we added a LeakyRelu layer after dense layer. Detail structure on this model are shown on table 2. For the hyperparameters, 100 epochs, 128 batch size is adopted and $adam$ optimizer is used.

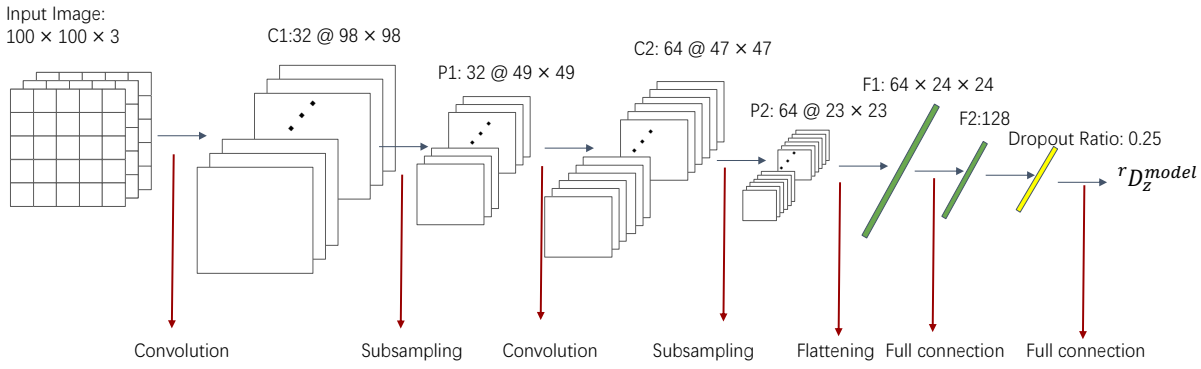


Figure 3: CNN Structure

Layer Type	Output Shape	Activation
Input	(100, 100, 5)	
Conv 2D(3×3)	(98, 98, 32)	ReLu
MaxPooling 2D(2×2)	(49, 49, 32)	
Conv 2D(3×3)	(47, 47, 64)	ReLu
Dropout(0.25)	(47, 47, 64)	
MaxPooling 2D(2×2)	(23, 23, 64)	
Flatten	36856	
Dense	128	linear
LeakyRelu	128	LeakyReLu(0.2)
Dropout(0.25)	128	
Dense	1	linear

Table 2: Structure of Multi layer CNN

4.2.3 Encoder Based Convolutional Neural Network (CNN)

Autoencoder (Ng et al. 2011) is an unsupervised machine learning technique to extract key features from input and ignore inessential point, learning the significant sets of data of input. Autoencoder consists of two parts, encoder and decoder, encoder responds to learning the crucial point from initial input, decoder is used to reduce refined features and return to its original state. In this model, we make use of an encoder, which is shown in below table 3, to perform feature extraction. The encoder is trained jointly with a ML structure as shown in table 3. Two different encoders were applied in this model, due to the different size of input image, vegetation map with a size of 100×100, and landform slope only has 30×30. Two encoder extracted feature from input separately and trained the concatenated dense layer. The detailed structure map are shown in figure 6 and table 3. In this study, the batch size, epochs and optimizer were set to 128, 20, $adam$, respectively.

Layer Type	Output Shape	Activation
EncoderA		
Input	(100, 100, 1)	
Conv 2D	(100, 100, 32)	ReLu
MaxPooling 2D(2×2)	(50, 50, 32)	
Conv 2D	(50, 50, 16)	ReLu
MaxPooling 2D(2×2)	(25, 25, 16)	
Conv 2D	(25, 25, 16)	ReLu
MaxPooling 2D(2×2)	(13, 13, 16)	
Flatten	2704	
LeakyRelu	2704	LeakyReLu(0.3)
Dense	169	linear
LeakyRelu	169	LeakyReLu(0.2)
Dense	20	linear
EncoderB		
Input	(30, 30, 1)	
Conv 2D	(30, 30, 32)	ReLu
MaxPooling 2D(2×2)	(15, 15, 32)	
Conv 2D	(15, 15, 16)	ReLu
MaxPooling 2D(2×2)	(8, 8, 16)	
Flatten	1024	
LeakyRelu	1024	LeakyReLu(0.3)
Dense	64	linear

LeakyRelu	64	LeakyReLu(0.2)
Dense	10	linear
Encoder Based CNN		
Input A	(100, 100, 1)	
Input B	(30, 30, 1)	
Encoder A (Input A)	20	
Encoder B (Input B)	10	
Concatenate	50	
Dense	50	ReLu
Dense	50	Linear

Table 3: Structure of Encoder Based CNN

4.2.4 Hyperparameter Tuning

To filter the optimal parameter value, grid search were applied on all models and compare resulting loss on test dataset. The number of estimator and max depth were studied for random forest. One more parameter, learning rate, was researched for XGBoost. Batch Size, epochs and optimizer were tested for both image-based method. Detailed result are shown on table 4.

Hyperparameters	Grid Search Space	Optimal Set
Random Forest		
Number of estimator	5, 10, 20, 50, 100	50
Max depth	1, 5, 10, 20, 50	5
KNN		
Number of neighbors k	5, 10, 20, 50	10
XGBoost		
Number of estimator	5, 10, 20, 50	50
Learning Rate	0.05, 0.1, 0.02	0.1
Max depth	1, 5, 10, 20, 50	5
Mutil-layer CNN		
Batch Size	32, 64, 128, 256	128
Epochs	10, 20, 50, 100	50
Optimizer	SGD, Adam	Adam
Encoder-based CNN		
Batch Size	32, 64, 128, 256	128
Epochs	10, 20, 50, 100	50
Optimizer	SGD, Adam	Adam

Table 4: Optimal Parameter of Grid Search

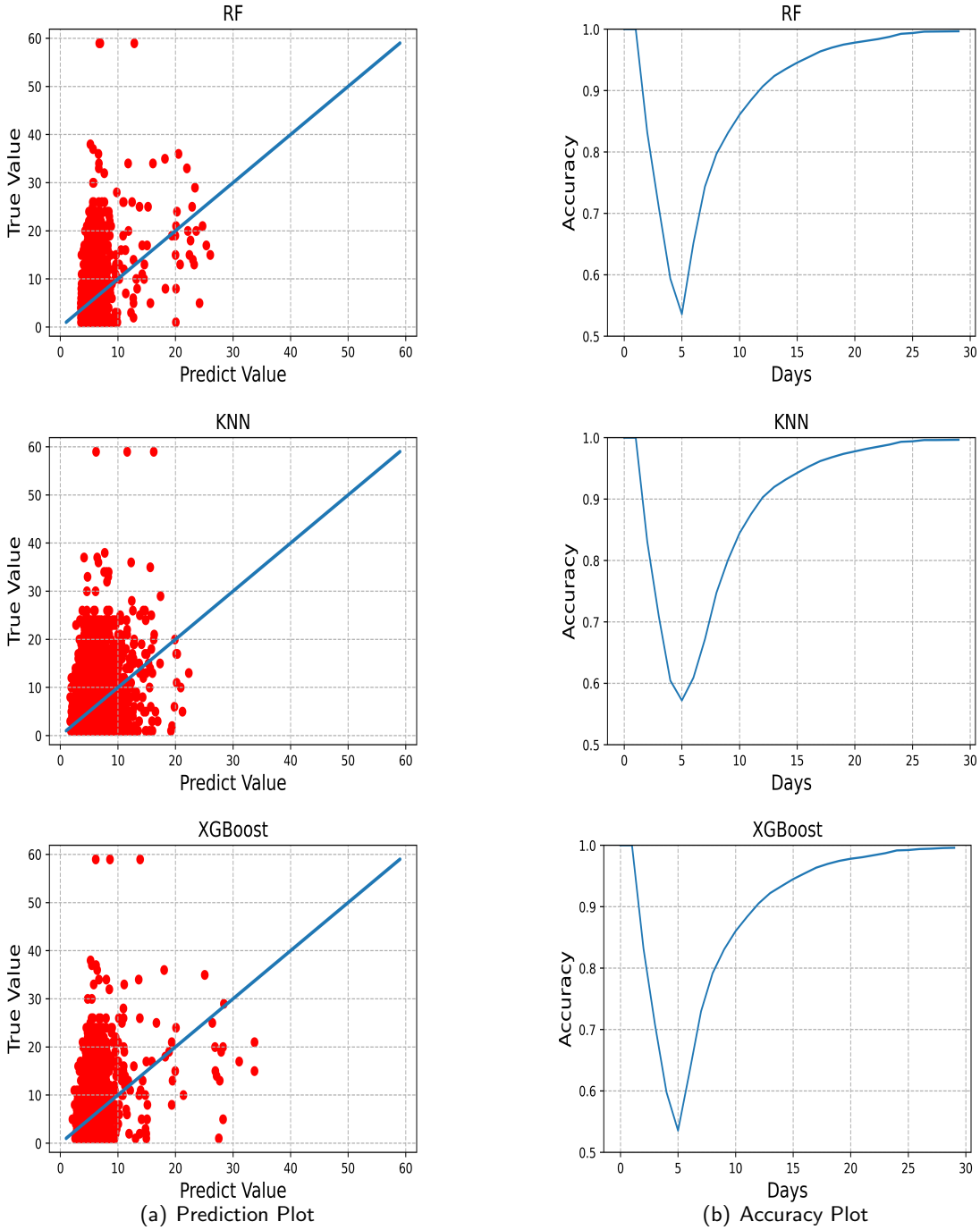
5 Results and Discussion

The importance of radius was proved unnecessary as Trend map shown on figure 8. Different Radius presents similar impact on RMSE value of prediction, thus radius is not a key factor to consider when new models come. All mentioned ML methods were used to predict the fire duration on the test set, fire from 2013 to 2016 on the target area, producing the prediction map for the results from all regression based ML approaches as shown in figure 4. Each plot has predicted value on x label and true value on y label and scatter point were used to indicate individual forecasting. The blue scatter line shown on the map represents the real trend of fires duration. Additionally, figure 4(b) give the accuracy plots that provide another measurement to identify the preciseness of models. The feature importance of RF and XGBoost were shown in figure 7. The Outcome Summary can be seen in Table 5.

Model	RMSE	R^2Score	Accuracy
Random Forest	4.99	0.073	0.536
KNN	5.26	-0.031	0.572
XGBoost	5.05	0.049	0.535
Multi-layer CNN	5.05	0.051	0.629
encoder-based CNN	5.16	0.005	0.631

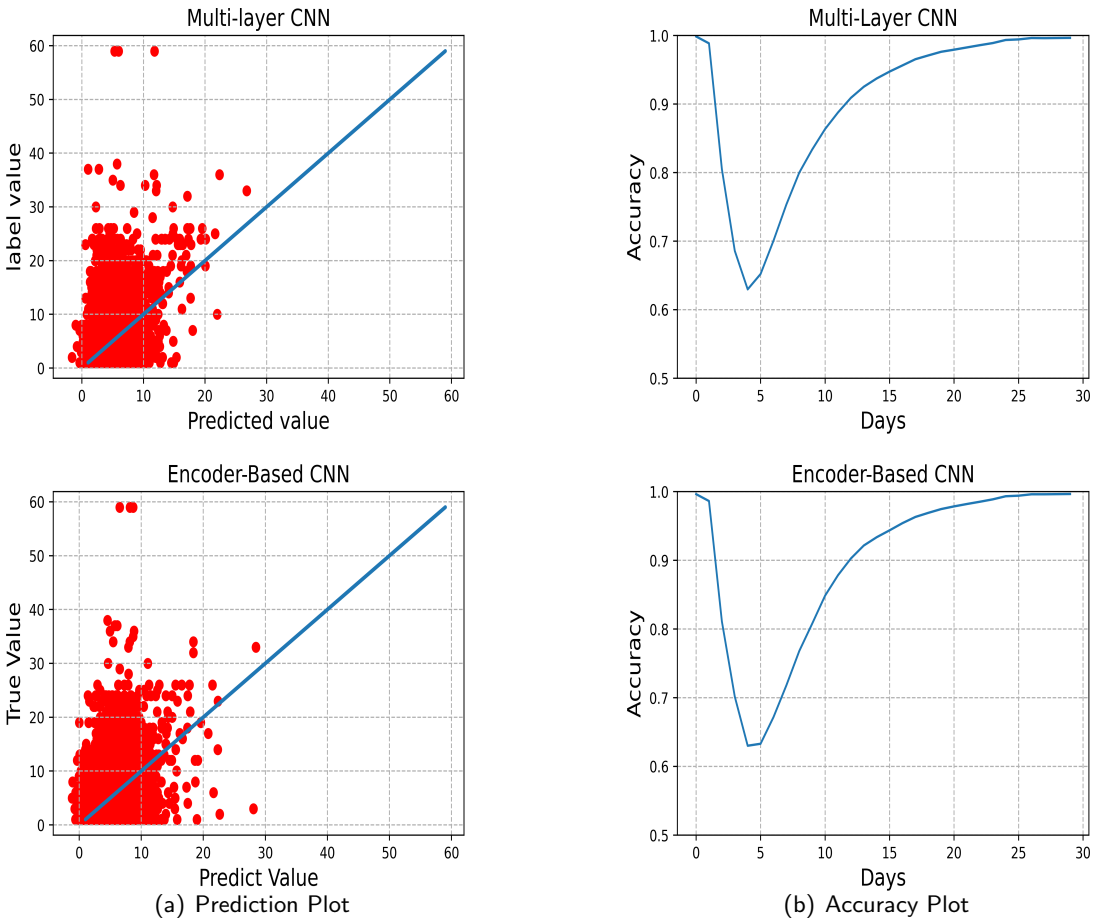
Table 5: Result Summary

The result shown in above Table 5 reveals that all applied ML methods presented no big different on RMSE, although random forest performs the best. Moreover, all the models lack a reasonable R^2 score for some reason, probably due to the large number of small fire, which incurs large data bias that not well quantified by R^2 score. Normally, a R^2 score that above 0.3 presents a reliable "goodness-of-fit" of model, but our results give R^2 score value all below 0.1, which indicates the inaccuracy of the models. However, some researchers argue that R^2 score may be inappropriate to evaluate the performance of a model and state that accuracy of a model does not only depend on R^2 score(Onyutha 2020). Thus, the accuracy rate is more valuable than R^2 score for referencing the reliability of model. In ML regression model, KNN performs the best accuracy, achieving a worst situation around 0.572. RF and XGBoost have nearly the same accuracy rate due to the similarity of these two method. For image-based models, below two figure 5(a) and 5(b) give the prediction and accuracy plot of two image based ML model. It is clear that image-based model demonstrate a more precisely output, which improves around 14 percent in accuracy. Although two model presents similar result, they still have distinctiveness. Encoder-Based CNN is more suitable for the scenarios that different size of input are involved, multi-layer CNN is applicable to features of the same size. All plots performed the accuracy close to 90% at days equal to 10 and approaching 100% at days moving to 30, indicating all the models are more sensitive to long-lasting fire and more precise in predicting large fire.



(a) Prediction Plot (b) Accuracy Plot

Figure 4: regression model result



(a) Prediction Plot (b) Accuracy Plot

Figure 5: regression model result

6 Conclusion

We developed a machine learning-based wildfire prediction model that learns from historical fire data and corresponding landscape maps to achieve fast predictions. The average error of the model can be controlled within one day and has above eighty percent accuracy for fires that last longer than ten days. The models have difficulties in predicting extreme cases, such as the prediction of fires of very long or short duration. Besides, our model cannot achieve a good R^2 score as R^2 score cannot evaluate result properly, another statistical coefficient is needed to replace R^2 score. This model is not suitable for the case that which small amounts of data and images were provided because our model requires amounts of input to learn a reasonable weight and bias. In addition, the input feature maps may change when applied to new target areas due to the feature importance of landscape features may change in a new region. Future enhancements to the model should focus on improving the regional coverage and generality and the simulation of features not considered in current situation.

Acknowledgments

I would like to thank Dr. Sibo Cheng for his help on helping all the problem in the three months, providing model establishing direction, giving crucial advice on model detail, and all the suggestion on writing report.

Main notations and acronyms

${}^{r}\mathbf{X}_{z}^{tc}$	Image of tree cover tc with fire index z in Radius r
${}^{r}\mathbf{X}_{z}^{gc}$	Image of grass cover gc with fire number z in Radius r
${}^{r}\mathbf{X}_{z}^{slope}$	Image of slope $slope$ with fire number z in Radius r
${}^{r}\mathbf{D}_{z}^{model}$	Predicted Duration of model with fire number z in Radius r
CNN	Convolution Neural Network
RF	Random Forest
KNN	K-Nearest Neighbors
$XGBoost$	eXtreme Gradient Boosting
CFD	Computational fluid dynamics

Appendix

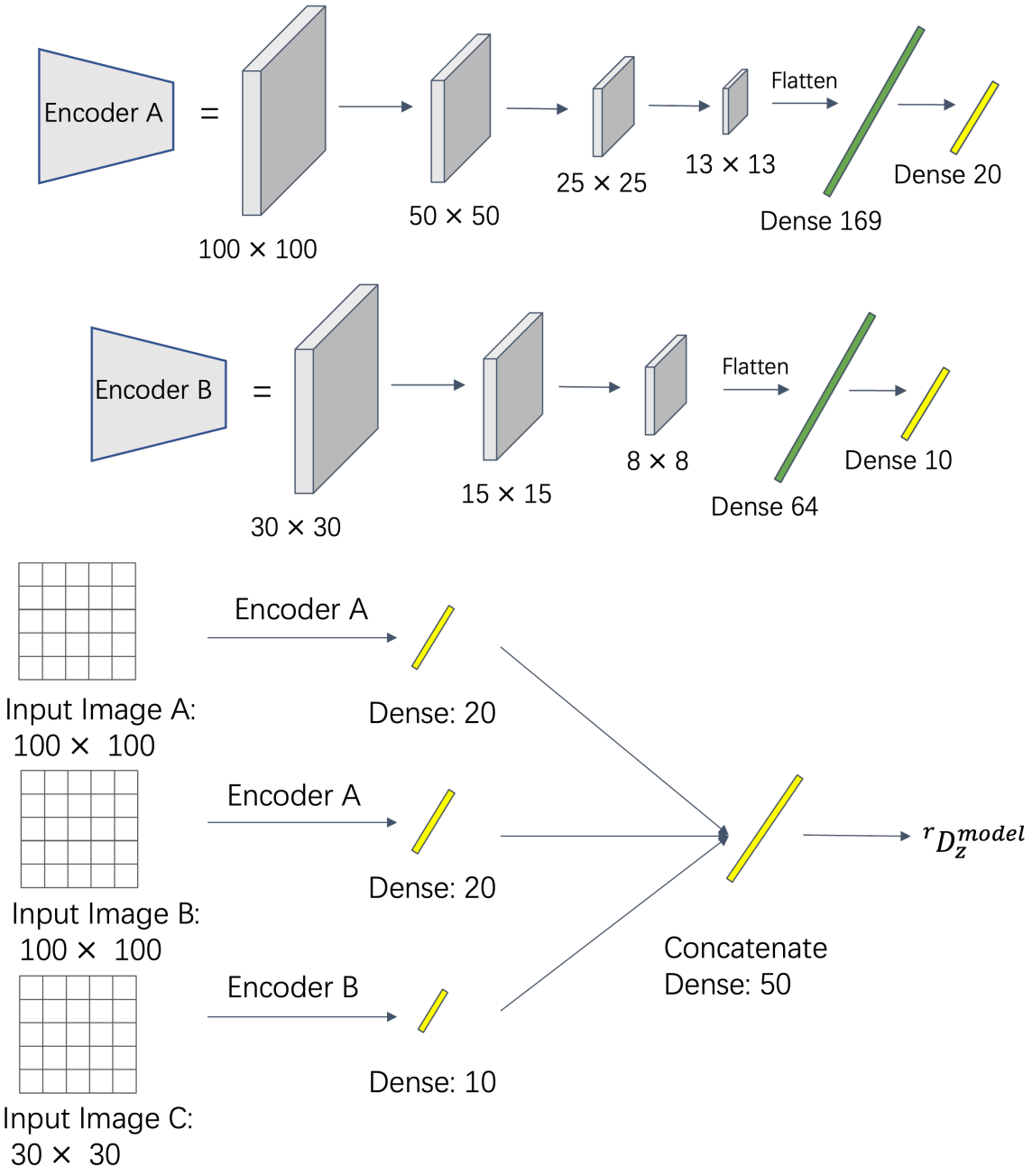


Figure 6: Encoder Based CNN Structure

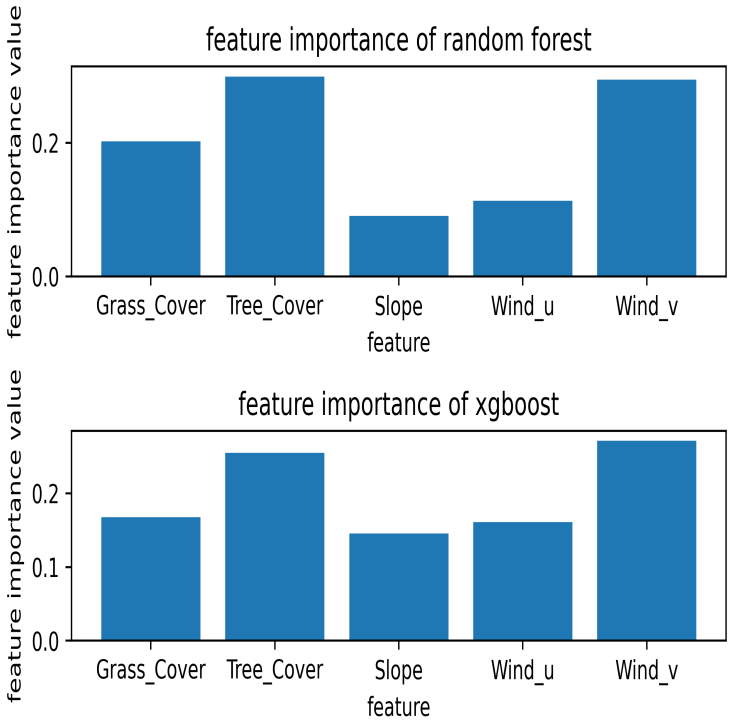


Figure 7: Feature importance Plot

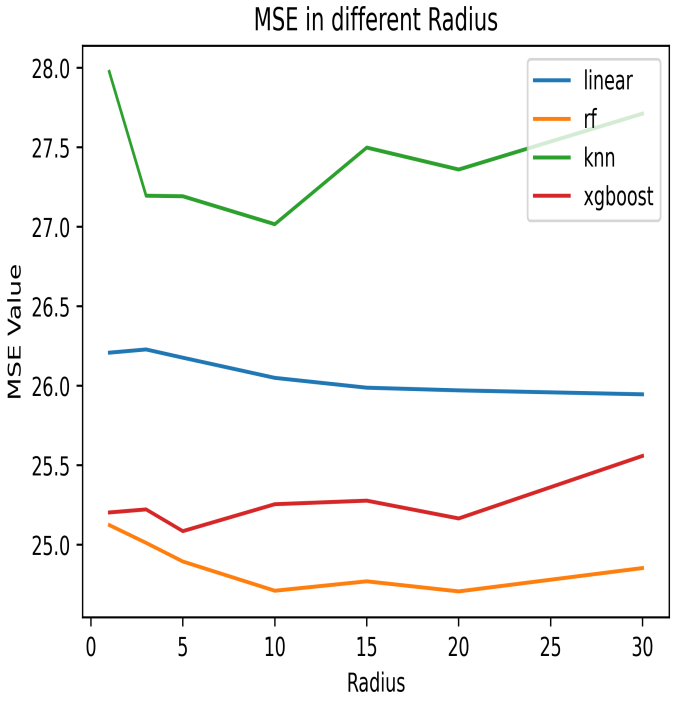


Figure 8: Caption

References

Albawi, S., Mohammed, T. A. & Al-Zawi, S. (2017), Understanding of a convolutional neural network, *in* '2017 international conference on engineering and technology (ICET)', Ieee, pp. 1–6.

Alexandridis, A., Vakalis, D., Siettos, C. I. & Bafas, G. V. (2008), 'A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990', *Applied Mathematics and Computation* **204**(1), 191–201.

Altman, N. S. (1992), 'An introduction to kernel and nearest-neighbor nonparametric regression', *The American Statistician* **46**(3), 175–185.

ANDELA, N., MORTON, D., GIGLIO, L. & RANDERSON, J. (2019), 'Global fire atlas with characteristics of individual fires, 2003-2016'.

Anderson, J. D. & Wendt, J. (1995), *Computational fluid dynamics*, Vol. 206, Springer.

Boesch, D. F., Field, J. C. & Scavia, D. (2000), *The potential consequences of climate variability and change on coastal areas and marine resources: Report of the Coastal Areas and Marine Resources Sector Team, US National Assessment of the Potential Consequences of Climate Variability and Change, US Global Change Research Program*, number 21, US Department of Commerce, National Oceanic and Atmospheric Administration ....

Bond, W. J. & Keeley, J. E. (2005), 'Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems', *Trends in Ecology Evolution* **20**(7), 387–394.

Breiman, L. (2001), 'Statistical modeling: The two cultures (with comments and a rejoinder by the author)', *Statistical science* **16**(3), 199–231.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L. & Smets, B. (2020), 'Copernicus global land cover layers—collection 2', *Remote Sensing* **12**(6).

Chen, T. & Guestrin, C. (2016), Xgboost: A scalable tree boosting system, *in* 'Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining', pp. 785–794.

Chopard, B. & Droz, M. (1998), 'Cellular automata', *Modelling of Physical* .

Dennison, P. E., Brewer, S. C., Arnold, J. D. & Moritz, M. A. (2014), 'Large wildfire trends in the western united states, 1984–2011', *Geophysical Research Letters* **41**(8), 2928–2933.

Fradkov, A. L. (2020), 'Early history of machine learning', *IFAC-PapersOnLine* **53**(2), 1385–1390.

Ghorbanzadeh, O., Valizadeh Kamran, K., Blaschke, T., Aryal, J., Naboureh, A., Einali, J. & Bian, J. (2019), 'Spatial prediction of wildfire susceptibility using field survey gps data and machine learning approaches', *Fire* **2**(3).

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I. et al. (2018), 'Era5 hourly data on single levels from 1979 to present', *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* **10**.

Jayalakshmi, T. & Santhakumaran, A. (2011), 'Statistical normalization and back propagation for classification', *International Journal of Computer Theory and Engineering* **3**(1), 1793–8201.

Kandakji, T., Gill, T. E. & Lee, J. A. (2020), 'Identifying and characterizing dust point sources in the southwestern united states using remote sensing and gis', *Geomorphology* **353**, 107019.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. & Liu, T.-Y. (2017), 'Lightgbm: A highly efficient gradient boosting decision tree', *Advances in neural information processing systems* **30**.

Madigan, J., Rowe, J., Angelos, J., Herthel, W., Matz, D., Dinucci, M. & Fletcher, V. (2011), '(a323) wildfire associated burn injury of 1400 sheep in northern california: A coordinated mass casualty veterinary response', *Prehospital and Disaster Medicine* **26**(S1), s90–s91.

Mastromichalakis, S. (2020), 'Alrelu: A different approach on leaky relu activation function to improve neural networks performance', *arXiv preprint arXiv:2012.07564* .

McCaffrey, S. (2004), 'Thinking of wildfire as a natural hazard', *Society and Natural Resources* **17**(6), 509–516.

Ng, A. et al. (2011), 'Sparse autoencoder', *CS294A Lecture notes* **72**(2011), 1–19.

Onyutha, C. (2020), 'From r-squared to coefficient of model accuracy for assessing" goodness-of-fits"', *Geoscientific Model Development Discussions* pp. 1–25.

Parisien, M.-A., Snetsinger, S., Greenberg, J. A., Nelson, C. R., Schoennagel, T., Dobrowski, S. Z. & Moritz, M. A. (2012), 'Spatial variability in wildfire probability across the western united states', *International Journal of Wildland Fire* **21**(4), 313–327.

Radke, D., Hessler, A. & Ellsworth, D. (2019), Firecast: Leveraging deep learning to predict wildfire spread., *in* 'IJCAI', pp. 4575–4581.

Robinne, F.-N., Bladon, K. D., Miller, C., Parisien, M.-A., Mathieu, J. & Flannigan, M. D. (2018), 'A spatial evaluation of global wildfire-water risks to human and natural systems', *Science of The Total Environment* **610-611**, 1193–1206.

Rodrigues, M. & de la Riva, J. (2014), 'An insight into machine-learning algorithms to model human-caused wildfire occurrence', *Environmental Modelling Software* **57**, 192–201.

Rothermel, R. C. (1972), *A mathematical model for predicting fire spread in wildland fuels*, Vol. 115, Intermountain Forest & Range Experiment Station, Forest Service, US ....

Smith, P. (1981), 'Bilinear interpolation of digital images', *Ultramicroscopy* **6**(2), 201–204.

Spawn, S. A., Sullivan, C. C., Lark, T. J. & Gibbs, H. K. (2020), 'Harmonized global maps of above and belowground biomass carbon density in the year 2010', *Scientific Data* **7**(1), 112.

Stetler, K. M., Venn, T. J. & Calkin, D. E. (2010), 'The effects of wildfire and environmental amenities on property values in northwest montana, usa', *Ecological Economics* **69**(11), 2233–2243. Special Section - Payments for Ecosystem Services: From Local to Global.

Theobald, D. M., Harrison-Atlas, D., Monahan, W. B. & Albano, C. M. (2015), 'Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning', *PLOS ONE* **10**, 1–17.

Youssouf, H., Liousse, C., Roblou, L., Assamoi, E., Salonen, R., Maesano, C., Banerjee, S. & Annesi-Maesano, I. (2014), 'Quantifying wildfires exposure for investigating health-related effects', *Atmospheric Environment* **97**, 239–251.